Segmentation of volumetric objects using the curve skeleton and watershed

Denis Horvat, IEEE Member

Abstract—This paper presents a new method for part-type segmentation of volumetric objects. The method iteratively merges segments from the object’s curve skeleton that represent stable geometry. Object segmentation is achieved by dilating the merged segments using the watershed algorithm, bounded by the inverted distance transform. Because the curve skeleton preserves the object’s general shape even at a presence of a surface noise, its influence on the segmentation is minimal. Additionally, the user can control the merging process by changing the merging criteria. Thus, different degrees of segmentation can be achieved that ranges from fine to coarse.

Index Terms—shape segmentation, shape representation, merging, curve skeleton, distance transform, watershed

I. INTRODUCTION

Object segmentation partitions 3D models into disjoint regions that are homogeneous with respect to one or multiple characteristics [1]. This is useful in data compression, geometric modelling, object manipulation, reverse engineering in CAD, shape analysis, shape retrieval, texture mapping, and collision detection [2, 3, 4, 5]. According to the literature [2], segmentation methods can be categorized as a surface-type and a part-type. The surface–type segmentation segments the object’s surface into surface patches, which are mainly useful in the computer graphics (e.g. texture mapping, remeshing, or simplification). The goal of the part-type segmentation, however, is to segment a shape into meaningful parts. For a given object, the number of parts and their extent is not formally defined and is, in many cases, application-dependent. However, segmentation should generally produce regions that a human intuitively perceives as a distinct part of the shape.

This paper presents a method for the part-type segmentation of volumetric objects, represented with a 3D grid (i.e. voxel models). The shape characterization relies on the objects’ curve skeleton. By defining the decomposition of the skeleton using key voxels, segments of the skeleton are identified, where the geometry of the object is considered as stable. Stable neighboring segments can then be iteratively merged based on the angle between them. The resulting segments are then used as seeds for the dilation with the watershed algorithm where inverse distance transform sets the dilation priority. Thus, the regions contained within the final segmentation meet at junctions with the smallest surface areas.

The advantages of the proposed approach can be outlined as follows: (1) the curve skeleton retains its general shape even when dealing with noisy objects and thus, the segmentation is still accurate; (2) the merging criterion of segments is controlled by the user which enables fine or coarse segmentation.

The paper is divided into 4 sections. The following subsection considers related works. Section II describes the proposed segmentation procedure, while results are given and discussed in Section III. Section IV concludes the paper and considers the possibilities of the method’s future improvements.

A. Related work

The majority of existing approaches perform the segmentation using scalar functions that are defined on the object’s surface. These include harmonic functions, piecewise harmonic functions, or eigenfunctions of the Laplace Beltrami operator [6, 3, 7]. Some approaches also consider pose-consistent segmentations, which identify segments consistently over several object transformations (e.g. different poses). The pose-consistent segmentation is achieved by the method proposed by Aubry et al. [8], which characterizes the relation of surface points at various spatial scales. Similarly, Huang et al. [4] achieved the pose-consistent segmentation using modal analysis, which enabled the identification of parts of the shape that tend to move rigidly. Several machine learning approaches have also been proposed. Xie et al. [9] suggested a segmentation using Extreme Learning Machine (ELM) classifier to define the initial segmentation, on which it then applies a graph-cut optimization constrained by the super-face boundaries. A constrained spectral learning approach is introduced by Sharma et al. [10] that is used to supervise the segmentation of a shape from a training data set, followed by a probabilistic label transfer algorithm to match two shapes and to transfer cluster labels from a training-shape to a test-shape. Svensson and Baja [11] proposed a method in which they group significant voxels that are found in relation to the distance transform. After region growing with the constrained inverse distance transformation, the obtained segments are merged based on an adjacency matrix. Hierarchical shape segmentation is proposed by Reiners and Telea [5], where the approach recognizes critical points on the curve skeleton and then calculates their component sets using geodesics.

This work was supported by the Slovenian Research Agency under Grants J2-6764 and P2-0041. Denis is with the Faculty of Electrical Engineering and Computer Science, Maribor, Slovenia (e-mail: denis.horvat@um.si).

ISBN: 978-0-9803267-8-9
II. METHODOLOGY

Let $E \subseteq \mathbb{N}^3$ represent a 3D binary grid, where each cell (i.e., voxel) $v \in E$ is defined by the coordinate triple $v = (x, y, z)$ and contains a value of either $0$ or $1$. Voxels with the value $1$ are referred to as white voxels $E^W \subseteq E$ and represent the object that will be segmented. The topology of $E^W$ is defined using the $3 \times 3 \times 3$ neighbourhood, where $N_v \subseteq E^W$ contains a set of neighboring white voxels to $v$.

The segmentation procedure proposed in this paper relies on the object’s curve skeleton. By observing the angle between the skeleton’s segments, geometrically different parts of $E^W$ can be identified. Segmentation of $E^W$ is then achieved by dilating the labeled segments, while conditioning the dilation priority with the distance to the nearest non-white voxel. Thus, the segments meet at junctions with the smallest surface areas, which are considered as borders between significant sections of an object. The procedure can be outlined using the following four steps:

1. **Curve skeleton calculation**: To efficiently characterize the object’s shape, its curve skeleton is calculated as proposed by Xie et al. [12].

2. **Segment initialization**: Stable segments of the skeleton are found, which represent part of the object, where the geometry is considered stable and should thus not be segmented.

3. **Segment merging**: The neighboring stable segments are iteratively merged, based on the number of voxels they contain and the angle between them. Non-essential segments are filtered.

4. **Segment dilation**: Using the processed segments as markers, a marker controlled watershed is used to obtain the final segmentation.

In the next subsections, each of the mentioned steps is described in detail.

A. Skeleton calculation

As several efficient approaches for the calculation of the curve skeleton already exist (see e.g. [13, 14, 15, 12]), we consider its construction beyond the scope of this paper. The curve skeleton $E^S \subseteq E^W$ is a compact descriptor that captures the shapes’ essential geometrical features by mapping them to a graph structure with univariate arcs [15, 16]. The curve skeleton is here obtained with the thinning algorithm proposed by Xie et al. [12], which uses sequential thinning iterations of $E^W$. Each thinning is characterized by six parallel directional sub-iterations, followed by a set of sequential sub-iterations. The algorithm is insensitive to object rotation and is only moderately sensitive to the noise. An example of an object and its curve skeleton that will be used as an example can be seen in Fig. 1.

While $E^S$ highlights the object’s geometric features, a procedure is needed to interpret them properly. The following subsections describe the two-step interpretation procedure of $E^S$, which identifies segments that represent objects within $E^W$ that should be segmented.

![Fig. 1: Object and its curve skeleton.](image)

B. Skeleton preprocessing

The interpretation procedure firstly identifies the stable segments of the curve skeleton. These represent sections of $E^W$ with no significant changes within the object’s geometry. Stable segments are defined using 3 types of key voxels, which are categorized depending on the number of their neighbors:

- **End voxels** are voxels where $|N_v| < 2$.
- **Regular voxels** are voxels where $|N_v| = 2$.
- **Crossroad voxels** are voxels where $|N_v| > 2$.

All 3 voxel types are shown in Fig. 2a, where the end voxels are marked with the red colour, the crossroad voxels are green, and the regular voxels are yellow.

![Fig. 2: Showing (a) key voxels of $E^S$ and (b) its merged segments based on $t^\Phi$.](image)

A set of neighboring regular voxels that are topologically located between any two non-regular voxels then form a stable segment, which also includes two key voxels that confine it. One region within the object can, however, be represented by multiple neighbouring stable segments, which should be merged in order to avoid oversegmentation. The exception to this are the singleton segments, which always represent only one region (e.g. the bottom left segment in Fig. 2). Singleton segments are stable segments confined by two end voxels, or voxels where $|N_v| = 0$. Thus, merging as described in the next subsection is only considered for non-singleton segments.
Merging of stable segments is done iteratively where a currently processed segment, denoted as $s_m$, always has a highest number of voxels. A successful merging step combines $s_m$ with one of its neighbour segments $s_i$. We say that $s_i$ is a neighbour segment to $s_m$ if both segments contain the same inroad voxel $v_{mi}$. The merging criterion is based on $\phi$, which represents the absolute angle difference from $180^\circ$ between a pair of vectors. These are defined by the three non-regular voxels; $v_{mi}, v_m, v_i$ as $v_m - v_{mi}$ and $v_i - v_{mi}$. $v_m$ and $v_i$ represent non-regular voxels in $s_m$ and $s_i$, respectively, where $v_{mi} \neq v_m \neq v_i$. If the smallest $\phi$, obtained between $s_m$ and $s_i$, is less than a user-defined threshold $t^\phi$, the segments are merged and $s_i$ cannot be merged with any other $s_m$. This merging step repeats until the merging criteria is satisfied (i.e. $s_i$ exists where $\phi < t^\phi$). Otherwise, $s_m$ is considered as complete and the process continues with merging other non-complete segments. To prevent oversegmentation, non-singleton segments that are smaller than 10 voxels, are removed. After filtering, we also remove the inroad voxels that still belong to multiple segments. Complete segments of the curve skeleton from Fig. 1 are shown in Fig. 2b. Using prioritized segment dilation, these remaining segments are then used to obtain the segmentation of $E^W$ as described next.

D. Segment dilation

Because the thickness of an individual object within $E^W$ can vary significantly, the dilation is done using the watershed algorithm [17, 18]. This is a region-based segmentation approach, where each region has a growing priority. Watershed can be illustrated using the following analogy [19]: A topographic relief with holes is immersed in a lake. Basins will fill up with water starting at the holes, and, at points where water coming from different basins would meet, dams are build that represent the edge of a region. In the proposed approach, the topographic map is obtained by the inverse distance transformation on $E^W$, while the remaining isolated segments of $E^S$ represent the holes. Thus, the regions meet at junctions of two objects with the smallest surface area. The watershed is performed until every voxel in $E^W$ has been marked with a segment label.

III. RESULTS

In order to test the proposed method, the segmentation was tested on voxelized polygonal meshes with different geometric characteristics. The voxelization was done using the binvox software [20, 21].

Fig. 3 shows the segmentation of the multiple models using the proposed approach. As seen, all parts of the voxel model, which can be considered as separate, are segmented accordingly. However, some errors are visible at the edges of some segments. This mainly happens because, during the execution of the watershed, unlabeled voxels within areas that belong to different segments can have an equal growing priority (i.e. equal distance to the nearest non-white voxel). Because of this, a given region can be over or under dilated. This, however, usually occurs with shapes that do not have clearly defined edges. For shapes such as the plane, the boundaries of individual segments are determined accurately.

![Fig. 3: Segmented test datasets.](image)

By changing the value of the merging criteria $t^\phi$, the proposed procedure also enables the user to control the degree of the segmentation. As shown in Fig 4, when $t^\phi$ is small, fewer segments of $E^S$ are merged and the objects within the model are segmented into multiple segments. Thus, by using the threshold $t^\phi = 30^\circ$, 7 segments are created (see Fig. 4a). On the other hand, when $t^\phi$ is increased, the segmentation is more coarse as are more segments from $E^S$ are merged. Thus,
when \( t^\phi = 40^\circ \), \( E^S \) is segmented into 6 segments and \( t^\phi = 50^\circ \) only 5, as seen in Fig. 4b and Fig. 4c, respectively.

To further test the robustness of the segmentation, high and low frequency noise was introduced to the surface of the used models prior to voxelization. Fig. 5 shows the dinosaur model affected by different noise patterns and its subsequent segmentations. When comparing the results of the segmentation achieved with and without the noise, we see that in cases shown in Fig. 5a and Fig. 5b the same objects were segmented from \( E^W \). This is to be expected because the method relies on the curve skeleton for its shape definition, which is not sensitive to noise. Thus, when noise is not strong (i.e. does not severely deform the object), new segments in \( E^S \) are not created and the skeleton of the object maintains its general shape. Stable segments are therefore still merged and the segmentation remains identical.

![Fig. 5: Segmentation of a dinosaur model with the added (a) high and (b) low frequency noise pattern.](image)

**REFERENCES**


